DOCUMENT RESUME

ED 346 109 TM 018 325

AUTHOR Guimberteau, Bernadette
TITLE Extracting Meaning from Physics Texts: Standards of Evaluation Used by Novices.
PUB DATE Apr 92
NOTE 47p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Processes; *College Students; Comparative Analysis; *Evaluation Criteria; High Achievement; Higher Education; Hypothesis Testing; Low Achievement; *Models; *Physics; Problem Solving; *Reading Comprehension; Student Attitudes
IDENTIFIERS *Comprehension Monitoring; Error Detection; *Evaluation Standards; Text Processing (Reading)

ABSTRACT

This research investigates comprehension monitoring by studying the relation between the perception of comprehension during reading and detection of comprehension failures. Two groups of four college students (good and poor learners) defined by a post-hoc median split learned a physics chapter while rating their comprehension on a four-point scale and justifying it after each sentence. It was found that poor learners process sentences at the local level and use fuzzy referents to talk about the text, while good learners are more exact, specific, and global. Poor learners detect many of their comprehension failures, but they pay more attention to what they understand than to what they do not understand, which leads them to label their comprehension as "approximate" instead of "poor" even for severe comprehension failures and to go on reading instead of trying to solve their problems. Poor learners are also deficient at evaluating hypotheses that they make to solve comprehension failures. Thus, the proposed new paradigm shows that detection of a comprehension problem does not necessarily mean good comprehension monitoring as is typically the case in the traditional error detection paradigm. Three tables present study data and there is a 32-item list of references. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED346109

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Extracting Meaning From Physics Texts: Standards of Evaluation Used by Novices.

Bernadette Guimberteau

University of California at Berkeley

AERA, 1992

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Bernadette Guimberteau

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 018 325

BEST COPY AVAILABLE

Abstract

This research investigates comprehension monitoring by studying the relation between perception of comprehension during reading and detection of comprehension failures. Two groups of four subjects (good and poor learners) defined by a post hoc median split learned a physics chapter while rating their comprehension on a four-point scale and justifying it after each sentence. It was found that poor learners process sentences at a local level, and use fuzzy referents to talk about the text whereas good learners are more exact, specific, and global. It was also found that poor learners do detect conceptual comprehension failures but they pay more attention to what they understand than to what they don't. This leads them to label their comprehension as "approximate" instead of "poor" even for severe comprehension failures and to go on reading instead of trying to solve their problem. Finally poor learners are deficient at evaluating hypotheses that they make to solve comprehension failures. I then discuss how these findings can provide new insight on several results of the comprehension monitoring literature.

Comprehension monitoring is a metacognitive skill that has received quite a lot of attention in the literature. It is commonly regarded as being made of two components: an evaluation phase during which the learner evaluates her current state of ongoing comprehension, and a regulation phase that comes into play when the learner has found her understanding inadequate (Baker & Brown, 1984).

A common way to operationalize the comprehension monitoring concept has been through the use of the error detection paradigm (Baker, 1979; Brown, Brandsford, Ferrara & Campione, 1983; Markman, 1979; Markman & Gorin, 1981; Otero & Campanario, 1990; Wagoner, 1983; Walczyk & Hall, 1989). In this research program, the subjects are typically asked to read texts containing problems such as contradictions. The fact that the problem is not detected is interpreted as an inadequate evaluation of one's comprehension. If the problem is detected, the subsequent behavior of the subject is examined. It is generaly found that overall subjects are poor at detecting comprehension failures (Baker, 1989). It has been suggested that this poor monitoring is due to the fact that poor readers do not use appropriate standards to evaluate their comprehension (Baker, 1985; Ryan, 1984). For instance, according to Baker, poor readers tend to use a lexical standard while reading.

Poor comprehension monitoring has also been observed in the research program related to the importance of examples in learning. The seminal paper of Chi, Bassok, Lewis, Reimann, & Glaser (1989) about the role that subjects' self-explanations can have when subjects are working on partially worked-out examples in physics has shown that poor learners and poor problem-solvers generate less monitoring statements than good students. Moreover, by distinguishing between monitoring statements that indicate understanding and those that indicate a failure to understand, the authors have also illustrated the fact that poor students detect comprehension failures less frequently than good students.

This result has been replicated by other studies in the same domain (Ferguson-Hessler, 1990), and in computer science as well (Pirolli & Bielaczyc, 1989; Pirolli & Recker, 1991). Pirolli and Recker, for example, had subjects learn about recursion. Analysis of the verbal protocols collected during the learning process showed that good subjects produced more monitoring utterances, and more metacognitive statements in general, when they processed the text as well as the examples. Thus, the correlation between the quantity and the nature of subjects' monitoring statements and the quality of effective understanding appears to be robust, and suggests a critical relationship between the two variables.

Pressley and his colleagues have recently shown that calibration of comprehension improves when the subjects are required to answer adjunct questions as they read and when these questions are similar to those presented on the postest (Pressley, Snyder, Levin, Murray & Ghatala, 1987). But in another experiment they showed that college students tend to have little awareness of the accuracy of their answers to thematic questions after having read short passages (Pressley, Ghatala, Woloshyn & Pirie, 1990). The subjects in this experiment read 20 short passages taken from practice tests for the Scholastic Aptitude Test. After reading each passage, they were asked a thematic question such as: " The message that the author wishes the reader to get is that..." In one condition, the students had to select their answer among multiple choice items. In another condition they were asked to generate a short answer. After answering, they were given a 7-point scale to rate how certain they were of their answer. Then they proceeded to the next passage. Results show that students in both conditions had significantly higher confidence ratings in correct than in incorrect answers, thus indicating some awareness of the quality of their understanding. But the students had a high confidence in the majority of the incorrect answers. It thus seems that adults often have a high confidence in the accuracy of their answers when these are

not correct. Pressley and his colleagues suggest that adults are not able to detect comprehension problems during reading.

What precedes gives a somewhat dark picture of subjects' comprehension monitoring abilities, aggravated by the fact that different operationalizations of the concept converge toward the same picture. There are, however, reasons to believe that the situation is not as bad as it seems. Winograd & Johnson (1982) have emphasized the fact that the error detection paradigm confounds several variables. Failures to detect anomalies in a text may be due to factors other than poor comprehension monitoring skills. Among these are: background knowledge, beliefs about the learning situation, and repair strategies that do not appear in the protocols. Readers may not detect a problem because they do not have enough background knowledge to identify it. They can also have a representation of the learning situation in which the learner is supposed to acquire true knowledge. They can also detect the anomaly but not report it because they rationalize it or they wait for more information to solve it. Thus, the poor readers of the error detection studies may actually turn out to monitor their understanding better than it is usually believed.

Subjects' rationalization about textual material is well illustrated by what happened in the study of Baker (1985). She asked college students to evaluate the coherence of a passage about Renaissance artists. This text included a paragraph about daVinci's drawings that had been designed to appear inconsistent. On the one hand, the text said that the drawings were prescient. On the other hand it stated that they lacked practicality. A large number of students did not report this inconsistency. They felt they had understood the passage because, as shown in interviews, they had rationalized the inconsistency the following way: they thought that the drawings were indeed prescient but from what they knew, it was true that they also lacked praticality because, for instance, the airplaines were aerodynamically unsound so they would never have been able to fly.

In another study, Otero and Campanario (1990) had 10th and 12th grade subjects study six texts supposedly from physics and chemistry textbooks with the purpose of evaluating their understandability. In the first phase of the experiment, subjects were given a booklet containing the texts, each presented on a different page. Some of them had been made so that they included an explicit contradiction involving a main point in the text. The students were asked to rate the comprehensibility of each text on a four-point scale, to underline conflicting sentences if they found any problems in the text and to explain the reasons for the problems. In a second part, the subjects were informed about the contradictory sentences and were asked, if they had not underlined the problematic sentences, why they had not done so. Among the subjects who detected the inconsistencies quite a few tried to fix them by building a coherent text base by activating some schema. The two schemata more frequently used were: "the advance of science" (to explain the contradiction by assigning the contradictory sentences to different developmental stages of science) and "there are no rules without exceptions".

What preceeds still tends to give a somewhat confusing portrait of a learner's ability to monitor her knowledge. On the one hand, learners could be better detectors of comprehension failures than previously believed. On the other hand, they can be confident in the correctness of incorrect answers on postests. They also may detect more comprehension failures than usually thought, but let us notice that if mechanisms such as rationalization occur frequently when the subject detects a contradiction in a text, the end product of the learning is not good monitoring anyway! Rationalization is likely to introduce a bias in the comprehension.

The above discrepancy may be due to the fact that research on comprehension monitoring has seldom paid attention to a critical factor of the monitoring process: the learner's perception of her own understanding while learning. I would like to point out that defining the evaluation phase of the comprehension monitoring phenomenon in terms of the perception of comprehension variable presents several advantages. First, the

concept emphasizes the subject's own definition of comprehension failure instead of an apriori definition in terms of the text characteristics or the meaning that the author intended to convey. For the error detection paradigm, a poor comprehension monitoring is defined by the fact that subjects do not detect textual problems during reading. For the perception of comprehension paradigm that I propose, poor monitoring occurs when: (a) a subject states she understands although it is clear that she does not, and (b) the subject believes that she does not understand, although she actually does. This latter aspect of poor comprehension monitoring has been, so far, somewhat neglected in the metacognitive literature. It is the other side of the metacognitive coin, but important, too, if one wants to carefully study how a subject's understanding is built and cognitively monitored.

Another advantage of the perception of comprehension concept is that it generates a more stringent test to detect poor monitoring than in the error detection paradigm. To be concrete, situations that would be labeled "good monitoring" in the error detection paradigm can turn out to be an example of poor monitoring in the perception of comprehension paradigm. For instance, a subject can detect a comprehension failure in a text but erroneously not believe that her comprehension is poor. This is a case of poor comprehension monitoring, but it would be counted as an instance of good monitoring in the error detection paradigm because the subject did detect the textual problem.

A third advantage is that it allows one to consider the dynamics of meaning-making from text in a more complete way. For instance, one can also study how subjects perceive their comprehension while they solve comprehension failures and how this perception affects the problem-solving process. In other words, the introduction of this concept into the comprehension monitoring framework allows us to consider that the regulation phase commonly proposed also has an evaluation phase of its own. This makes it possible for the model to take into account and try to explain new facts that have so far been neglected. Interesting cases of poor monitoring strategy can also occur when a

subject perceives she has not solved a comprehension failure although she has and vice versa. Such situations are also relevant to understand how subjects build up scientific reasoning skills (Guimberteau, 1991).

Another appealing advantage to using the perception of comprehension concept is that it allows the study of the perception of poor understanding as well as the perception of good understanding. The error detection paradigm tries to provide answers as to why a subjects does not understand. Its purpose is not to focus on why the subject thinks she understands. One could say: if poor learnes tend to detect problems at the lexical level may be they think they understand on the basis of a lexical rationale. But there maybe more than that. By focusing on the reasons why subjects think they understand, one can obtain more clues as to why learners do not report detected comprehension failures when this lack of report is related to a feeling of understanding.

Focusing on the perception of good as well as poor understanding has another advantage. It allows one to study how a subject balances what she perceives she does not understand with what she perceives she does. It may be that subjects pay more attention to what they understand than to what they don't. For instance, when reading a sentence, if they understand part of it but detect a comprehension failure at the lexical level, they might still think that their understanding is good overall. This could explain the discrepancy obtained between data coming from verbal protocols and behavioral responses to contradictions. Although contradictions are not necessarily reported verbally, measures such as galvanic skin responses show that the GSR recordings can fluctuate sharply when a subject is reading the contradictory information (Tikhomirov & Klochko, 1981).

Another situation that can occur is that a subject can rate her comprehension as poor as soon as she detects a comprehension failure, for instance if she does not understand one part of a sentence even though she understands another part. Still, another scenario is that the subject can use some form of "gravity scale" for the

comprehension failures. If the comprehension failure involves a concept the problem can be regarded as severe whereas if it does not, the subject will see her understanding as good anyway.

It may also be that, in some contexts, treatment of what is not understood depends upon what is understood. That is, in some cases, the subject can recognize a comprehension failure whereas she will not in other cases. For instance, if the subject has the feeling to understand much of what has preceded, maybe she will pay less attention to a comprehension failure than if she has the feeling of not understanding what precedes.

The purpose of the present study is to investigate the following issues: What are the standards that subjects use to evaluate their comprehension as good when they do not detect a comprehension failure? To which extent do subjects modify their comprehension when they detect a comprehension problem? Especially: can they perceive their comprehension as overall good in some cases and overall poor in others? On which basis do they do so?

Method

Subjects

Twelve subjects who volunteered for the experiment were run in this study. They had either no background in physics, or one high-school physics course. They were paid $5.00 per hour for their participation. Male and female students were equally represented in the sample.

Materials

The chapters used in this research were written by the author of this study from a standard textbook in introductory physics (Haber-Schaim, Cross, Dodge & Walter,

1971). They presented motion in a straight line, vectors, and motion in two dimensions. The chapters included few formulae, and especially no calculus, except for the concept of "limit" that was graphically used to explain the ideas of instantaneous velocity and acceleration.

Design and Procedure

Subjects were individually tested. They were told that the experiment was about the study of cognitive processes involved in the learning of scientific materials. They were allowed to use any study method that they like, such as taking notes, or highlighting the material.

Subjects began the experiment by some practice about verbal protocols. The goal of this first step was to have them practice verbalizing (1) while learning from a text, and (2) while solving problems. Two tasks were used. The first task consisted in talking aloud while learning a two-paragraph text about electricity. The second task required the subject to answer three questions (give the definition of "velocity" and "instantaneous acceleration", define the difference between speed and velocity). The purpose of this second task was also to screen from the subjects' answers the individuals that were already sophisticated in physics from those who were not. The former were not run in the experiment.

After going through this screening process, the subjects learned, in two successive sessions, chapter one on motion in one dimension and chapter two on vectors and were tested after each of these chapters. In a third session, a new chapter "Motion in two dimensions" was introduced. Half of the subjects were told that they would be given a test on this chapter when they finished learning it, and the other half was told that they would have to write down notes at the end of the session about what they remembered[2].

After learning the chapter, all the subjects were tested. Then, the subjects in the "notes" condition wrote down the notes, and the subjects in the "test" condition were also asked to write down notes. The data obtained in this part of the experiment have not been used in the present paper.

A few days later, subjects were assigned to the same "notes" ot "test" condition, but were told that the focus of the last session was different than the focus of the previous ones. In this new session, researchers were interested in finding more about why people think they understand or not. They were told that they would learn the chapter about "Motion in two dimensions" according to the following procedure. After each sentence, they would score their level of understanding on a 4-point scale made of four little boxes labeled "1" to "4" at the top. They were told to check "1' when they understood quite well what they had just read, "2" if they felt they understood only approximately, "3" if they did not understand, and "4" if they did not really know how well they understood. They were also asked to give detailed reasons to explain their rating. The instructions ended by stressing the fact that there were many reasons underlying feelings of understanding, and that the researchers in this study were interested in any kind of reason subjects in this study could possibly experience. The subjects were then asked to take some time to memorize the scale and they were given a one and a half page text on electricity to practice this procedure. They were then handed out the chapter. The text and the figures were the same as the ones presented in the third session. There were, however, several differences regarding the layout. In the first version, the text was double spaced whereas it was triple spaced in the new version to leave room for the comprehension scale after each sentence. Each line on the page did not necessarily have the same number of words as the corresponding line in the first version because space was used for the scale. Finally, each page did not necessarily include the same information as in the first version because the space taken by the scale after each sentence sometimes made it physically impossible to make a given page ends

exactly at the level its corresponding page in the first version ended. After learning the chapter, subjects took a test made of 17 problems: 9 isomorphic, and 8 transfer. Then, they all wrote down notes about what they remembered.

Results

The results below contrast two groups respectively called "good learners" and "poor learners". These groups were defined after the third session, on the basis of the subjects' performance to the postest after learning the corresponding chapter. Only the top-scoring four and bottom-scoring four students have been used in the analysis. The respective average performances of these two groups were 91 and 25 % correct. Each of these group is made of two subjects from the "notes" condition and two subjects from the "test" condition. The manipulation of the purpose of study turned out not to be significant.

Each of the several systems of categories that have been used in the analysis have been evaluated for reliability by having two coders code 60 statements (10 per subject) coming from 3 poor learners and 3 good learners, pooled and randomly presented. A given system of classification was used if the correlation between the codes was at least .80.

Comprehension rating

The sentences that the subjects read could be evaluated by choosing among four possibilities: "1" when the subject thought she understood the sentence well, "2" or "3" when the understanding was respectively approximate or poor, and "4" when the subject was undecided. How did the subjects do? Table 1 below gives a general overview of the rating distribution.

Insert table 1 about here

This table gives two types of information. First, it shows that the good learners perceived that they understood the material more often than the poor subjects (a Pearson RxC Chi Square with corrections gives $z = 5.9627$, $p < .01$). This means that there is some global sense of awareness in the two groups of what is actually going on in terms of comprehension.

Second, the poor subjects' scoring indicate that these subjects tended to qualify their understanding as "approximate" more often than "poor" in situations of comprehension failure. Of the total number of comprehension problems detected by the learners, only 28% were qualified as relating to a poor understanding. An hypothesis that could explain this behavior would be that these subjects drew a distinction between different types of comprehension failures. I shall investigate this later.

The above table does not tell us about the standards used by the subjects to decide about the quality of their understanding. What follows is thus a finer-grained investigation of the evaluation procedure used by the subjects.

Perception of good comprehension in good and poor learners

To find out about the existence of non-detected comprehension failures, the protocols for sentences that have been coded "good understanding" have been analyzed according to three types of standards of evaluation. The mention of the train of thought, the partial or complete processing of the ideas contained in a sentence, and the specificity of the referents used by the subjects when giving their justification of the scoring.

In terms of the train of thought, justifications for scoring good understanding have been assigned to a "global" category when the subject was able to relate the meaning

of the sentence to the train of thought (that is, to the goal of the authors, to the means to attain these goals, or to differences between different parts of the chapter). Two extracts below provide illustrations of this situation.

Sentence:

One way of finding the instantaneous velocity in this case is to note that the horizontal component of the velocity is constant.

A:

I understand this very well ... hmm... we are still working with the case of uniform horizontal velocity, they are restating what we are working on here.

Sentence:

Because the average acceleration is constant throughout the motion, we can be almost certain that the instantaneous acceleration is also constant throughout the motion.

A:

OK I understand this very well, I can see the connection that they are making here, average acceleration, constant horizontal velocity, and then instantaneous acceleration is almost certainly constant. I can see how this follows from the initial assumptions.

By opposition, a subject's response was classified as "local" when she did not relate the meaning of the sentence to the train of thought.

Sentence:

At the same time the ball on the right covered the same vertical distance between each image as did the ball on the left which had no horizontal velocity.

S:

I can see what they mean on the graph there are equal points along the vertical line at each interval OK so I think I understand that pretty well....

Sentence:

We can rewrite the last equation so that it more closely resembles the equations we developped for the description of motion along the straight-line path.

S:

OK I believe that you can probably rewrite that.

Analysis of the justifications classified this way shows that good learners tended to be more global than the poor ones ($z = 3.604$, $p < .01$). Good learners tend to be global in their way of processing sentences from a chapter, that is, they keep the "general picture" in mind and relate sentences to this picture. In contrast, poor learners are local. For instance, they have rarely mentioned the givens of the situation (constant vertical acceleration and constant horizontal velocity) when learning the first part of the chapter. This result is coherent with the fact that good learning from a text involves the building of a situation model to represent the situation in the text (Van Dijk et al. 1983).

Besides the above analysis in terms of train of thought, the verbal protocols were analysed a second time in reference to the total or partial processing of each sentence, as revealed by the verbalizations. The categories that emerged as a coding scheme are presented in the appendix.

It was found that there is a strong relationship between learner type and the nature of sentence processing as revealed by the protocols. Good learners tend to give a justification that relates to a complete processing of the sentence whereas poor learners

are associated with a partial processing ($z = 4.109$, $p < .01$). Good learners' justifications tend to be explanatory in nature, such as the ones below:

Sentence:

Let us analyze its motion by finding its instantaneous velocity at successive intervals as it falls.

R:

OK again I'm checking off 1 I realize that they want us to analyze the motion by checking the instantaneous velocity, that is, trying to figure out, what is the speed at any given point in time, at any given point along that path how fast is the ball moving that's how we're going to check out its motion.

Sentence:

The average acceleration is therefore constant.

R:

OK if you take the acceleration from over an interval that's called average acceleration, this average does not change... average acceleration... that is the average change in velocity is the same.

These examples show a systematic strategy to explain sentences: the reformulation of the sentence by coming back to the definition of the concepts. This is a strategy that has been commonly used by the good learners of this study.

In addition to the previous coding schemes, the verbal protocols were also coded in terms of the nature of the referents used by the subjects. This criterion concerns the use of demonstratives ("that") or pronouns ("it"). Subjects were coded "specific" when they could exactly refer to the idea of the sentence. Example:

Sentence:

Chapter 3: two-dimensional motion.

S:

I check "1" because I can picture what two dimensional motion means in my mind.

In contrast, subjects were coded "fuzzy" when they used demonstratives or pronouns. Examples :

Sentence:

Now let us examine what happens to the motion of an object that moves horizontally with uniform velocity but is also accelerated vertically (motion in two dimensions).

S:

I understand that... hmm.. because I can picture that in my mind... we're examining what happens in that case and I can see that case.

Sentence:

Figure 3.4 shows that the instantaneous velocity vector of the ball changes steadily

S:

I understand what they mean and I can see it on the graph so that gets a "1".

Sentence:

Let us analyze its motion by finding its instantaneous velocity at successive intervals as it falls.

E:

OK I understand that so I check 1...If we're going to do it at successive intervals it's not going to be complicated, it's going to be successive intervals and that's how we're going to find it.

It was found that good subjects tend to be more specific in their answers than poor subjects ($z = 8.209$, $p < .01$). The fact that they can also give fuzzy answers to a justification does not decrease their tendency to be specific. Many of their fuzzy responses include answers that they had clearly stated previously and that they did not want to repeat. This is not the case for poor subjects.

The above analysis was carried out to understand in which conditions subjects can rate their understanding as good without detecting a comprehension failure that does exist. It has been shown that such situations can exist because of the tendency of poor learners not to reason at the global level, not to process sentences totally, as well as their satisfaction with fuzzy referents instead of specific ones.

Perception of comprehension problems in poor learners.

Detection of comprehension failures. One of the critical questions that was raised in the introduction of this paper was the following one: do subjects detect comprehension failures or are they oblivious to them as it has typically been written in the comprehension monitoring literature? In what follows, I am going to discuss the problem of conceptual failures because the text was built around several concepts (e.g. "velocity", "instantaneous acceleration") that were critical for a good understanding. In addition, we will examine to which extent the poor learners of this study detected comprehension failures related to the critical points of motion in a curved path.

The verbal protocols show that poor learners mentioned all the concepts they had problems with in a straightforward manner (e.g. "... I am not positive of... exactly what

'uniform horizontal velocity' means...", "I'm not sure what the 'average acceleration' means..", "...I just have this thing with 'instantaneous velocity", "I don't have a clear definition of what 'instantaneous velocity' is). In addition, the subjects stated some problems more than once. For example, S. mentioned her problem with "uniform velocity" at three successive intervals ("hmm... I'm not really sure what 'uniform velocity' means", "I am still not very clear on the 'uniform horizontal velocity", "... hmm... 'uniform velocity').

However, the conceptual comprehension failures were not necessarily identified right away. For instance, D. read the sentence "Now let us examine what happens to the motion of an object that moves horizontally with uniform velocity but is also accelerated vertically". She checked her understanding as "poor" because "it's long and I just have to repicture the objects so I'll read it again". The second time, she rated "2" on the comprehension scale. "I understand that better... I can picture what they mean...".

If the poor learners of this study detected their conceptual comprehension failures, did they try to solve them? If so, how? S., E. and J. used several strategies. They sometimes stated their comprehension problem but went on reading, probably hoping to clarify the meaning of the concept by further reading. At other times, they tried to solve their problem by analyzing the sentence in which it was occurring, by looking back at the definition of the concept in a previous chapter, or by using visual features to come up with a meaning.

Besides D. who would underline the dificulties but rarely tried to solve them., the other subjects exhibited a similar pattern. Typically, when they would try to solve a problem, and propose an hypothesis, they would be deficient in the last step, the evaluation of the hypothesis. For instance, S. either tested her hypothesis by rereading the sentence or, more often, she simply did not test it and concluded on the basis of her feeling if the hypothesis was right or wrong. It means that sometimes she regarded a solution as correct although it was not (she believed that uniform velocity applied to the

vertical dimension). She could also reject a hypothesis that was correct (she rejected "steady" as a substitute for "uniform"). When she thought that her hypothesis was wrong, she usually went on reading instead of trying to propose a new one. E. detected a comprehension failure, with "instantaneous velocity" , tried to solve it by finding the definition of the concept. She did not fully understand the definition so she came back to the last chapter to reread the sentence that had initially caused a problem. She did not try to use the definition in this sentence. She went on reading.

Besides this problem-solving pattern, the protocols show that S. and D. were deficient in the ability to accept givens. They believed that everything had to be explained. Thus, they did not know how to explain the fact that the right ball in figure 1 has an initial uniform velocity. This is not to be explained, it is simply part of the givens of the situation that is analyzed.

There were six critical ideas necessary to achieve a good understanding of motion in a curved path. (1)The first one involved understanding the idea of "uniform horizontal velocty". (2) The second one was about the decomposition of an instantaneous velocity vector, one component being equal to the initial horizontal velocity vector. (3) Then, the subject had to understand the application of this method to each flash of the falling ball. (4) This led to observe a constant change of the instantaneous velocity vector. (5) The instantaneous velocity could then be expressed in terms of the initial velocity . (6) And, finally, the instantenous velocity vector could also be expressed in terms of the constant acceleration. Did the poor learners of the present study detect comprehension failures related to these ideas? Analysis of the protocols and of the teaching notes show that they did, most of the time.

Insert table 2 about here

In addition, it was observed that when subjects detected comprehension problems (that is, when subjects checked "2" or '3" on the understanding scale) they mentioned significantly more "strong" standards of evaluation such as context than when they rated their comprehension as quite good ($z = 1.98$, $P < .05$). An example of the use of context in situation of detection of comprehension failure is given below:

Sentence:

In each successive intervals, it (the instantaneous velocity vector) increases by the same amount, Δv.

D:

I can't see how it changes, how we know that it changes by the same amount, unless it says that it changes by 2 meters per second each time.

Here, D. refers to a piece of information that occurred two sentences ago.

Sentence:

The average acceleration is therefore constant.

S:

OK again, I understand that approximately hmm... it is the "therefore" which I am not entirely sure about so I am not sure about the sentence either.

In this example, S. does not refer to previously stated knowledge but express a concern for not being able to make a connection.

This observation of references to context in situation of poor or approximate understanding is somewhat counterintuitive. We have seen that poor learners tend to evaluate good comprehension on the basis of partial processing of a sentence. A criterion based on context implies more demanding processing than partial processing.

Thus, if these learners are able to partially process the above sentences, why don't they rate "good comprehension?" The fact that they do not means that subjects are not as poor detectors of comprehension problems than it is usually written. Poor learners can function with demanding standards to evaluate comprehension problems.

To conclude, the protocols show clearly that the poor learners of this study were able to detect most of their conceptual comprehension failures. They also used context to identify comprehension problems. However, they were deficient in the regulation phase of the monitoring process.

Discrimination between comprehension failures. We have previously seen that poor subjects tend to qualify a comprehension failure more often as "approximate" than "poor". Does this distinction reflect a difference in terms of the nature of comprehension problems that are detected? In other words, do poor subjects feel that their understanding is poor when they detect a critical comprehension failure and do they think it is approximate when the comprehension failure is minor according to their own standards? If this is so, what are these standards used to evaluate the importance of a comprehension failure? What follows is aimed at exploring these issues.

Three subjects in the analyzed sample of poor learners have used "3" ratings (poor comprehension): S., D., and E. Analysis of their justifications show that they score "3" when they cannot extract any meaning at all from the sentence for two reasons. First, this happens when the sentence does not mean anything the first time the subjects reads it (category 1). An example of this situation is offered by S. who reads "The position and velocity of the ball are shown here on a single graph" and says "hmm, I don't understand that, it just did not mean anything the first time I read it". The second situation occurs when the subject has tried to extract some meaning from the sentence but faces a comprehension failure (category 2). For example, D. reads "Therefore as

time progresses, v will point more and more downward" and says " I don't understand why that would happen." This category accounts for respectively 36%, 44%, and 33% of S., D. and E.'s ratings.

Did these subjects coded "2" rather than "3" for comprehension failures that are somewhat minor? In addition, did J., who used only "2" ratings, did so because of minor comprehension failures? To answer this question, I counted the occurrence of severe comprehension failures falling under the "approximate understanding" category. I defined a comprehension failure as "severe" when it fell under one of the two categories below:

1. The subject has a conceptual failure. Examples:

Sentence:

It was released at the same time as the ball on the left, but had a uniform horizontal velocity of 2.0 m/sec.

D:

I don't fully understand that sentence just because I am not positive of ... exactly what uniform horizontal velocity means.

Sentence:

What is the average acceleration in the time interval $\Delta t = t_1 - t_2$?

J:

allright, I'll put "2" 'cause I'm not sure what "the average acceleration" means but I can understand substracting one time interval from the other.

2. The subject does not understand the real point of the sentence. Example:

Sentence:

The average acceleration is therefore constant.

S:

OK, again, I understand that approximately, hmm... it is the "therefore" which is something I am not entirely sure about, so I am not entirely sure about the sentence either.

It turns out that the poor learners having used "2" and "3" ratings generated an average of 70% of severe comprehension failures under the label "approximate understanding". J.'s average was 78%. The results are given in the table below.

Insert table 3 about here

If the subjects did not make the distinction between "2" and "3" on the basis of the severity of their comprehension failures, what made them sometimes decide that their understanding was approximate and at other times that it was poor? It seems that the fact that they were able to understand part of a sentence was a triggering condition to label the understanding "approximate". Poor subjects never mentioned any form of understanding when they checked "3" whereas they did when they rated "2" ($z = 4.76$, $p < .01$).

Thus, it seems that poor learners rate "2" when they understand part of the sentence but detect a comprehension failure. However, they seem to be more influenced by what they understand than by what they don't in evaluating their understanding

because they do not really pay attention to the severity of the comprehension failure that they have detected.

What is it exactly that they understand when they rate "approximate understanding"? The subjects' justifications show two types of understanding: one relates to the logic of adjacent sentences and the other one concerns internal clauses. Examples of these types of understanding are given below:

Example of adjacent logic (coherence between adjacent sentences):

Sentence:

(Previous sentence is: In each successive interval it increases by the same amount Δv.")
The average acceleration is therefore constant.

J:

hmm I'll put a 2 again because I understand if in each successive interval it increases by the same amount that would lead to constant average acceleration however I'm still... some information...
I'm not sure about the complete picture

J., S., D. and E. respectively generated 53%, 23%, 20%, and 0% of justifications related to adjacent logic.

Example of internal clause:

Sentence:

If, instead of using 10-second intervals, we had used 2-second intervals, we would have gotten the pictures shown in Fig.3.7(a).

S:

hmm, I understand half of that sentence that there are using 2 second interval
I don't understand why they got the picture.

J., S., D. and E. respectively justified 47%, 77%, 80%, and 100% of their justifications with internal clauses.

Does the perception of understanding as approximate or poor influence the treatment that the not well understood sentence can receive? I made the hypothesis that a subject would try to solve a comprehension failure more often when she thought that her understanding was poor rather than approximate. It turns out that there is a category of comprehension failures related to a poor understanding that does not receive attention from the poor learners. This happens when equations occur. "I don't understand that... Ohhhhh (S. gives a start) what does it mean? ... hmm.. and I am not going to understand it more by continuing to read it (she turns the page)". In this experiment, comprehension failures associated with equations have typicallly been rated "3" by the poor learners with very few attempts to solve them.

To compare the "2" and "3" rated sentences, I also discarded from the "3" category sentences for which the subjects have not been able to make sense at all at first reading. These sentences are systematicallly reread. Finally, the comparison of comprehension failures labeled "2" with their "3" counterparts reveals a difference in terms of attempts to solve the problems. Subjects attempt to solve "3" rated sentences more often than the "2" rated ones ($z = 5.198$, $p < .01$).

In conclusion, it appears that poor learners do not distinguish between an approximate and a poor understanding on the basis of the severity of the comprehension failure that they have detected. To score an approximate understanding, they seem to pay more attention to what they feel they understand rather than to what they don't. This feeling of understanding makes them overestimate the quality of their comprehension in the presence of severe comprehension failures such as conceptual problems. They choose to qualify this understanding as "approximate" rather than "poor". This

perception is correlated with the treatment that they give to the sentence: they tend to go on reading after an "approximate sentence", probably hoping to solve the problem later; They stop and try to work on the problem in the case of "poor sentences", except if the comprehension problem involves an equation.

One could argue that subjects tend to focus on what they feel they understand because they have been pre-exposed to the material they had to learn in a previous session. In other words, they could think they understand because they have already seen the sentences. This interpretation is unlikely because a common observation in this experiment was that the poor learners seemed to process the sentences as if they were discovering them for the first time. For instance, they would express surprise when discovering the text ("Oh! It's a formula!") or rate their comprehension as poor and say "that did not mean anything". In the section about motion in a curved path they even read the following sentence: "What will be the path of such an object?" and said "I don't know" or "I can't picture the path although I understand the question" without ever mentioning that the path should be curved!

It seems that the above situation is the result of several factors. First, by definition, the poor learners of this study did not extract a lot of information from the chapter the first time they went through it. One can not expect them to recognize facts that they have not processed from the text. Another reason, related to the first one, is due to task demands. Subjects have been exposed twice to the text but in very different conditions. The first time, subjects were simply asked to read and think aloud while learning from a double-spaced text. The second time, they have to pay attention to each sentence and to stop after each sentence to rate their understanding. They are led to process the text at a more detailed level than they did during the spontaneous reading session. They discover meaning and wrap the text in a new identity. This can be illustrated by S. The first time she read the chapter, she built her interpretation around the idea that the uniform velocity and the constant acceleration mentioned in the

text were both in the vertical dimension. Her interpretation changed after the second reading. Acceleration was assigned to the horizontal direction..

The instructions are also different. Instead of simply being asked to talk aloud, the second session on the chapter emphasizes that the experimenters are interested in the reasons underlying feeling of understanding, whatever they may be. Subjects learn from a text that looks different from the first version because it includes a series of four little numbered boxes after each sentence. They check the box corresponding to their feeling of understanding, and explain the reasons for their scoring.

The above differences between the two situations probably explain why subjects mentioned rarely the fact that they had been through the chapter in the previous session. Such mentions were observed typically at the beginning of the session when subjects read the title of the subsection "Motion in a curved path". It also sometimes happened later in the chapter when the subject was working on a comprehension failure. S. mentioned the previous reading three times in the course of a comprehension failure. E. did so twice, J. and D. never mentioned it.

Discussion

In this experiment I explored the perception of comprehension and the standards of evaluation of comprehension used by poor and good learners during acquisition of knowledge. It was found that subjects, whether they are good or poor learners, have a pertinent global sense of their level of comprehension. Good learners tend to perceive their understanding as good. Poor learners use "good comprehension" ratings less often than good learners and have a considerably larger proportion of approximate and poor ratings.

Analysis of the standards of evaluation of good comprehension showed that good learners tend to give justifications of their scoring that relate to the train of thought, that is, the global picture of the chapter. They tend to process all the information

contained in sentences. They tend to give explanatory justifications rather than only descriptive. They are very specific in their use of referents, that is, in naming.

Poor learners tend not to learn in connection with the train of thought of the chapter, they process sentences only partially and in a more descriptive way than the goor learners. Finally, they use fuzzy referents to talk about the content of sentences.

A study of the justifications given by the poor learners when they rated their understanding as "approximate" shows that they tend to do so when they understand part of the sentence but experience a comprehension failure. Poor subjects use this type of rating even when they detect severe comprehension failures, such as conceptual failures or the inability to identify the real point of a sentence. This suggests that these subjects evaluate their comprehension more in relation to what they understand than in reference to their comprehension problems. Poor learners consider their understanding as "approximate" when they understand something related to the sentence and they state a poor comprehension when they do not understand anything. They tend to reread sentences that they have poorly understood and they try to solve their difficulty, (except if the sentence is an equation) whereas they prefer to go on reading if they think they understand approximately.

To explain what makes a learner rate her comprehension of a sentence as good although it is not, I have tried to understand the kinds of standards of evaluation that are used. It appears that such a phenomenon is correlated with the combination of the following factors: reasoning at the level of the sentence instead of the global level, and partial processing of the sentence. In addition, this phenomenon was also correlated with the use of fuzzy referents in the protocols.

The preference for a local processing and a local reasoning over the more global story presented in the chapter agrees with the fact that poor readers are said to be deficient at building situation models during reading (VanDijk & Kintsch, 1983). Several intervention programs have been implemented to help learners overcome this

deficiency. Subjects are given instructions for self-regulated reading that aim at teaching them how to better monitor their understanding (Palincsar & Brown, 1989). These instructions typically emphasize how to pay attention to the structure of the text while reading, how to generate questions for oneself and how to summarize. These techniques have proved efficient with fifth and sixth graders especially.

The fact that the poor learners of this study tended to use fuzzy referents in their protocols speaks for the necessity of designing intervention strategies for better reading that not only stress how to read better but also emphasize how to talk about a given text. Being able to talk about a text is not necessarily a sign of good comprehension. There are different ways to do so. It has been shown here that poor and good learners differ dramatically in terms of the exactness of their comments.

We have seen that the poor learners of this study pay more attention to what they understand than to what they don't when they detect a comprehension failure. On the one hand, they are better detectors of comprehension failures than it is usually believed. They are able to tell when their grasp of a given concept is not satisfactory. However, these poor learners focus too much on the meaning they are able to extract when they detect a comprehension failure, that is, on clauses within a sentence and, to a lesser extent, on the coherence of ideas between adjacent sentences. This fragmentary understanding leads them to overestimate the quality of their comprehension. They can believe that their understanding is approximate rather than poor in presence of a severe comprehension failure. They can also believe that their understanding is good although they have detected a comprehension failure. In addition, the fact that they perceive their understanding as approximate or poor is correlated with the absence or presence of attempt to solve their comprehension problem.

The above finding has several implications. First, this exacerbated sensitivity for understood knowledge adds a new insight on the causes of comprehension failures. Researchers have speculated on the origins of these problems. Baker and Brown

(1984), for instance, have identified four types of comprehension failures. First, the reader's insufficient knowledge about a topic can make the interpretation of a text problematic. Second, the author can not convey her thought in a clear manner which results in comprehension problems for the reader. Third, the reader can interpret the text in a coherent manner, but her interpretation can be different than the one the author intended to convey. Finally, the reader can construct a representation of the text similar to the one the author had in mind, but not be able to go beyond it and to critically evalua[illegible]. The present research proposes a new clue as to why comprehension failures occur. By paying too much attention to what she effectively understands in a sentence, the reader forgets to take care of gaps she has detected. She generates comprehension failures by not taking care of the ones she has detected.

A second important implication of the above finding lies in the fact that it contradicts the view according to which poor readers have little awareness that, as readers, they must attempt to make sense of the text. Subjects of low reading ability are usually depicted as decoders rather than sense makers (Baker & Brown, 1984; Paris, Wasik & Turner, 1991). They are described as focused on being able to say the words correctly. I have shown here the opposite. The poor readers in this study have too much a tendency to produce meaning. They are meaning oversensitive. In other words, they produce meaning for circumstances that good readers do not even bother to consider. This observation goes along well with the finding according to which less skilled comprehenders develop too many mental substructures during comprehension (Gernsbacher, Varner & Faust, 1990).

The above result also sheds some light on findings such as the one from the self-explanation effect literature according to which poor subjects tend to generate more monitoring statements indicating that they understand than indicating a comprehension failure while they learn (Chi et al. 1989; Fergusson-Hessler et al. 1990; Pirolli et al. 1991). It is generally suggested that poor learners mistakenly say that they

understand because they are not able to detect their comprehension failures. According to the result of the present study, the poor learners of the self-explanation effect literature may generate more statements indicating that they understand because they may actually pay more attention to what they have understood during learning than to the comprehension failures that they have detected.

The previous result also contradicts the conclusion of Pressley et al. (1990), which is that adults are not able to detect comprehension problems during reading. The authors made this conclusion after having noticed that subjects reading SAT passages and answering questions had a high confidence in the majority of the incorrect answers, although their confidence ratings for correct answers were higher than for incorrect answers. The fact that the subjects had confidence in incorrect postest answers does not necessarily mean that they did not detect comprehension problems as they were reading each text. The present research suggests that they did detect comprehension problems but that they built their comprehension of the text on what they thought they understood. This could explain why subjects in the Pressley et al. study were able to significantly discriminate between their correct and their incorrect responses.

Finally, the importance that poor learners give to what they understand when they detect a comprehension failure has essential implications for intervention measures. I believe that subjects should be taught about this tendency to foster their comprehension monitoring strategies especially in front of difficult scientific texts. Programs have, so far, been designed to teach subjects how to use better standards to evaluate their comprehension. Subjects are typically taught how to evaluate their comprehension at the levels of lexicality, external consistency, propositional and structural cohesiveness, internal consistency and informational completeness. Such training seems to lead to generalization in the use of standards and to a general improvement of performance (Baker & Zimlin, 1989). But these results have been obtained on a population of children, and in specific contexts of textual inconsistencies.

For example, application of an internal consistency standard is detected when a subject is able to tell that there is a contradiction in the following sentences: "The morning light was very bright. Chuck and Kenny wanted to go outside and play. Their mother said no because it was too dark." In a difficult, technical text, however, detecting contradictions is not always an easy matter, especially when there are no obvious problems coming from the text. How can the subject judge if some information is contradictory or not? How can the subject know that her understanding and the meaning in the text are contradictory? The issue here is that the knowledge of appropriate standards of evaluation for better comprehension monitoring is not necessarily enough for good monitoring because these standards are not always easy to implement for the learner. In particularly difficult texts such as some science texts, a learner may know about the issue of appropriate standards but not be able to use them appropriately and regress to the tendency of focusing on what she understands. Thus, in situations where a learner does not really know how to implement good standards of evaluation, the subject should be taught not to over-rely on comprehension. One way to do so would be to provide these subjects with the actual protocol of a poor learner, and explain how the learning cannot proceed because of too much emphasis on what the knowlegde the learner is able to extract.

Why did subjects pay too much attention to what they understood since they had quite often detected a comprehension failure? This is very likely due to the fact that they did not know how to solve their comprehension failure. Thus, in the present context, it seems that subjects would have benefited in standards to solve their comprehension failures and also in standards to evaluate a solution to a given comprehension failure. For instance, if S. had reread the first paragraph of the chapter after she had decided that uniform velocity applied to the vertical dimension, maybe she would have noticed that the text only mentions "uniform horizontal velocity".

Finally, this research aims at balancing the image given by the literature of misconceptions in science, especially in physics, to explain why learning science can be so difficult. It has been shown that naive subjects hold a lot of conceptions about the physical world that interfere with the acquisition of real scientific knowledge (Viennot, 1979; Trowbridge & McDermott, 1980; Clement, 1982; McCloskey, 1983). An example of such conceptions is the straight-down belief. If a man is walking at a constant speed and suddenly drops a ball that he was holding in his hand, then the ball will fall straight-down. The present research has attempted to demonstrate that learning physics and science in general is also difficult because a learner can make costly comprehension monitoring mistakes.

References

Baker, L. (1979). Comprehension Monitoring: Identifying and Coping With Text Confusions. Journal of Reading Behavior, 11, 363-374.

Baker, L. (1985). How Do We Know When We Don't Understand? Standards for Evaluating Text Comprehension. In D.L. Forrest-Pressley, G.E. Mackinnon & T.G. Waller, Eds., Metacognition, Cognition and Human Performance. New York: Academic Press.

Baker, L. (1989). Metacognition, Comprehension Monitoring, and the Adult Reader. Educational Psychology Review, vol. 1, No. 1, 3-38.

Baker, L. & Brown, A.L. (1984). Metacognitive Skills and Reading. In P.D. Pearson, M. Kamil, R. Barr & P. Mosenthal (Eds). Handbook of Reading Research, vol.1, pp. 353-394. White Plains: NY: Longman.

Baker, L. & Zimlin, B. (1989). Instructional Effects on Children's Use of Two Levels of Standards For Evaluating Their Comprehension. Journal of Educational Psychology, 3, 34-346.

Brown, A. (1988). Motivation to Learn and Understand: on Taking Charge of One's Own Learning. Cognition and Instruction, 5, 311-321.

Brown, A.L., Brandsford, J.D., Ferrara, R.A. & Campione, J.C. (1983). Learning, Remembering, and Understanding. In J.Flavell and E.M. Markman (Eds), Handbook of Child Psychology (4th ed.). Cognitive Development (Vol. 3, pp 515- 529). New York: Wiley.

Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P. & Glaser, R. (1989). Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems, Cognitive Science, 13, 145-182.

Clement, J. (1982). Students' Preconceptions in Introductory Mechanics. American Journal of Physics, 50, 66-71.

Ferguson-Hessler, M.G.M. & de Jong, T. (1990). Studying Physics Texts: Differences in Study Processes Between Good and Poor Solvers. Cognition and Instruction, 7, 41-54.

Gernsbacher, M.A., Varner, K.R. & Faust, M.E. (1990). Investigating Differences in General Comprehension Skills. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 3, 430-445.

Guimberteau, B. (1991). Cognitive Psychology and the Study of Scientific Discovery. Theory & Psychology, 2, 223-244.

Haber-Schaim, U., Cross, J. B., Dodge, J.H. & Walter, J. A. (1971). Physics. Lexington, MA: D.C. Heath & Co.

Markman, E.M. (1979). Realizing That You Don't Understand: Elementary School Children's Awareness of Inconsistencies. Child Development, 50, 643-655.

Markman, E.M. & Gorin, L. (1981). Children's Ability to Adjust Their Standards for Evaluating Comprehension. Journal of Educational Psychology, 73, 320-325.

McCloskey, M. (1983). Intuitive Physics. Scientific American, 4, 122-130.

Otero, J.C. & Campanario, J.M. (1990). Comprehension Evaluation and Regulation in Learning From Science Texts. Journal of Research in Science Teaching, 27, 5, 447-460.

Palincsar, A.S & Brown, A.L. (1989). Instruction for Self-Regulated reading. In L. Resnick & L.E. Klopfer (Eds), Cognitive research in subject matter learning (pp. 19-39). The Association for Supervision and Curriculum Development.

Paris, S.G., Wasik, B. A. & Turner, J.C. (1991). The Development of Strategic readers. In R. Barr, M. Kamil, P. Mosenthal & P.D. Pearson (Eds). Handbook of Reading Research vol.2, pp. 609-640. White Plains: NY: Longman.

Pirolli, P. & Bielaczyc, K. (1989). Empirical Analyses of Self-explanation and Transfer in Learning to Program. Proceedings of the 11th Annual Conference of the Cognitive Science Society, Hillsdale, NJ: Erlbaum.

Pirolli, P. & Recker, M. (1991). Knowledge Construction and Transfer Using an Intelligent Tutoring System. Technical Report No CSM-1, University of California at Berkeley, School of Education.

Pressley, M, Snyder, B.L., Levin, J.R., Murray, H.G. & Ghatala, E.S. (1987). Perceived Readiness for Examination Performance (PREP) Produced by Initial Reading of Text and Text Containing Adjunct Questions. Reading Research Quaterly, 22, 219-236.

Pressley, M. & Ghatala, E.S. (1990a). Self-Regulated Learning: Monitoring Learning from Text. Educational Psychologist, 1, 19-33.

Presley, M., Ghatala, E.S. , Woloshyn, V. & Pirie, J. (1990b). Sometimes Adults Miss the Big Ideas and Do not Realize it: Confidence in Responses to Short-Answer and Multiple-Choice Comprehension Questions. Reading Research Quarterly, 3, 232-249.

Ryan, M.P. (1984). Monitoring Text Comprehension: Individual Differences in Epistemological Standards. Journal of Educational Psychology, 76, 248-258.

Tikhomirov, O.K. & Klochko, V.E. (1981). The Detection of a Contradiction as the Stage of Problem Formation. In Wertsch, J. (ed.), The Concept of Activity in Soviet Psychology, Sharpe, Armonk, New York, pp. 341-382.

Trowbridge, D.E. & McDermott, L.C. (1980). Investigation of student understanding of the concept of velocity in one dimension. American Journal of Physics, 48, 12, 1020-1028.

Van Dijk, T. A. & Kintsch, W. (1983). Strategies of Discourse Comprehension. New York: Academic Press.

Viennot, L. (1979). Spontaneous Reasoning in Elementary Dynamics. European Journal of Science Education, 1, 205-221

Wagoner, S.A. (1983). Comprehension Monitoring: What It is and What We Know About It. Reading Research Quaterly, 18, 328-346.

Walczyk, J.J. & Hall, V.C. (1989). Is the Failure to Monitor Comprehension an Instance of Cognitive Impulsivity? Journal of Educational Psychology, 3, 294-298.

Winograd, P. & Johnson, P. (1982). Comprehension Monitoring and the Error Detection Paradigm. Journal of Reading Behavior, 14, 61-76.

Appendix

Coding scheme for the level of processing of a sentence:

1. States understanding without explaining.

Partial processing of the sentence

2 . rephrasing. Example:

Sentence: To find out we must determine the vector change $\Delta v = v2 - v1$ in the velocity.

J: OK, they are finding the difference and the vector change in the velocity. (subject does not process one part of the sentence: "to find out" and repeats the other part).

3 . explanation. Example:

Sentence: To confirm this, of course, we might want to take additional strobe photographs of the falling ball, using shorter and shorter time intervals to approach the limit as $\Delta t \to 0$.

J: I'll put a "1" for that sentence because I understand just by making a photograph of smaller and smaller time interval, shorter and shorter makes it easier to measure and closer to no differences at all between each time. (subject does not process "to confirm this", but explains what $\Delta t \to 0$ means).

Complete processing of the sentence

4 . rephrasing. Example:

Sentence: Another important type of two-dimensional motion is motion in a circular path.

J: OK, I'll understand that motion in a circular path is an important two-dimensional motion.

5. explanation. Example:

Sentence: Then its speed v is given by the circumference divided by the period: $v = 2\pi R/T$.

J: I think I understand... the velocity v is given by the distance it covers over the time.

6. Other. No extraction of words from the sentence, understanding is simply stated

. because the explanation was given before

. because of background knowledge

. description, "I see"

Author notes

Work reported in this paper was supported in part by a grant from the University of California at Berkeley and by Dr. Peter Pirolli (NSF 1-442120-21900-3).

I gratefully acknowledge suggestions from Peter Pirolli and Ann Brown as well as comments from Kate Bielaczyc, Mimi Recker, and Dan Berger in the design phase of this study.

Correspondance should be addressed to Bernadette Guimberteau, School of Education, EMST, University of California at Berkeley, Berkeley, CA 94720.

Footnotes

The subjects in this last condition were told that their notes would be given to subjects taking an experimental course. This was done to study the difference between learning to take a test versus learning to teach. The results turned out not to be significant.

Table 1

Percentage of Final Levels of Perception of Sentence Comprehension

	Good learners				Poor Learners			
	R.	D.	A.	C.	E.	J.	D.	S.
Good Compr.	98	97	95	96	80	60	61	53
Approximate Compr.	2	3	5	4	16	39	21	28
Poor Compr.	0	0	0	0	4	1	18	19
Undefined Compr.	0	0	0	0	0	0	0	0

Note. This table gives only the final level of comprehension chosen for each sentence. However, subjects did not necessarily rate their comprehension of a given sentence only once.

Table 2

Detection of Comprehension Failures Related to the Main Ideas of Motion in a Curved Path.

	Poor Learners			
Main Idea	E.	J.	D.	S.
(1)	N	-	Y	Y
(2)	Y	Y	Y	Y
(3)	Y	N	N	Y
(4)	-	Y	Y	Y
(5)	Y	Y	Y	Y
(6)	Y	Y	Y	Y

Table 3

Presence of Severe Comprehension Failures in Sentences Perceived as Approximately Understood

Poor Learner	% of Severe Comprehension Failures
D. (N= 26)	65
S. (N= 35)	69
E. (N= 21)	76
J. (N= 51)	78

Bernadette Guimberteau
School of Education
EMST
University of California at Berkeley
Berkeley, CA 94720

bg@garnet.berkeley.edu

Extracting Meaning From Physics Texts:

Standards of Evaluation Used by Novices.

Poster presented at AERA, San Francisco, 1992.

This research proposes a new paradigm to study comprehension monitoring. It investigates comprehension monitoring by studying the relation between perception of comprehension during reading and detection of comprehension failures. Two groups of four subjects (good and poor learners) defined by a post hoc median split learned a physics chapter while rating their comprehension on a four-point scale and justifying it after each sentence. It was found that poor learners do detect many of their comprehension failures but they pay more attention to what they understand than to what they don't. This leads them to label their comprehension as "approximate" instead of "poor" even for severe comprehension failures and to go on reading instead of trying to solve their problem. Thus, the proposed new paradigm shows that detection of a comprehension problem does not necessarily mean good comprehension monitoring as it is typically the case in the traditional error detection paradigm.